Model Error Estimation for SSSCs Delivered in Phase-1 by the Group2 Consortium

Joydeep Bhattacharyya[1], Nikolai Shapiro[2], Michael Ritzwoller[2], Hans Israelsson[1], Xiaoping Yang[1] and Keith McLaughlin[1]

1 SAIC, Arlington, Virginia
2 UC, Boulder, Colorado

Group2 Consortium Documentation, Phase-1 Delivery

September 2001

1.0 INTRODUCTION

The Group2 consortium uses travel time predictions of regional phases from 3-D mantle models used to relocate seismic events. Therefore, errors in the predicted travel times can directly affect the robustness in event locations. These errors in the travel time predictions are called *modeling errors* which combines, among other factors, with the errors in the arrival times of picked phases (*measurement error*), misidentification of seismic phases and errors in the origin error to give us uncertainty in event locations. In the location algorithm used in the Group2 Phase 1 delivery (Yang et al., 2001), independent estimates of the model error and the measurement errors are used to estimate confidence ellipses. The *a priori* model errors are a phase-dependent function of distance and control how the different phases are weighted in the relocations. Therefore, with realistic model error estimates, we also expect to get better locations since the data are weighted more appropriately in the inversion. In this report, we present the methodology used for developing and testing model error estimates, which form an integral part of the SSSCs (Nagy, 1996) used in the relocations.

The primary goal of the Group2 consortium is to develop regional traveltime tables that improve event location estimates in the Middle East, Mediterranean region and North Africa. This framework comprises model based Source Specific Station Corrections, SSSCs, that contain estimates of predicted travel time as a function of source location and their associated errors. In Phase 1 of this project, we developed SSSCs for Pn and Sn, for two different 3-dimensional Earth models, CUB1.0 and SAIC-HRV. Two related but separate sets of error estimates are provided for Pn and Sn. These model error estimates are the same for both the CUB1.0 and SAIC-HRV models.

The goals of this sub-project are the following:

a. Estimate model errors that captures un-modeled 3D structure;

b. Test if the error estimates predict "honest" coverage ellipses, i.e., does 90% confidence ellipses cover the Ground Truth location 90% of the time;

c. Propose improvements to methodology.

We present in this report a collection of analyses that support the choice of simple distance dependent model errors for the Phase 1 delivery. These include:

a. Re-analysis of the Pn and Sn misfit from 3° - 20° for the IASP91 travel times from the EHBGT5 catalog, i.e., the events from the EHB catalog (Engdahl et al., 1998) which are candidates for being considered to be of GT5 quality.

b. Analysis of Pn misfit from 3° - 20° of 1,00,000 Pn phases from the EHB catalog for the CUB1.0 3D velocity model. This allows us to ascertain the level of unmodeled signal in the travel time predictions of the data.

c. Breakdown of Pn misfit at IMS stations for the CUB1.0 model by geographic regions that shows us the spatial distribution of this unmodeled signal.

d. Comparison of empirical (JHD) path corrections with CUB1.0 and SAIC-HRV model predictions for event clusters allows us to directly test the models in certain regions.

e. Analysis of residual variance between empirical travel times (JHD) and 3-D model predicted path corrections that can help us separate modeling and measurement errors.

f. Estimation of the underlying variance of the Pn and Sn travel times using data from the IDC Reviewed Event Bulletin (REB).

g. A comparison of the travel time predictions between the CUB1.0 and SAIC-HRV models. The difference between the predictions can give us an estimate of modeling errors.

h. Analysis of coverage from relocation experiments using GT0 – GT10 reference events which allows us to test for "honest" coverage ellipses.

The error estimation methodology used in Phase 1 is simple and based on empirical data. Error estimates will be improved in Phase 2. Our goal in this analysis is to obtain error estimates that give statistically significant location improvements and "honest" coverage ellipses. In the following sections, we describe the modeling error estimates and justify their use in the Phase 1 location validation tests. Results from relocation experiments and cluster analysis, described in other sections of the Phase 1 delivery, have also been incorporated in this analysis. Finally, based on these experiments, we suggest possible methodologies that can be used to improve the modeling error estimates in the Phase 2 of the Group2 calibration project.

2.0 Estimation of Model Errors

We have constructed model error surfaces based on empirical information. These estimates were developed for both Pn and Sn phases. We have not developed a theoretical foundation of the error problem for this preliminary delivery. Estimates of the error propagation through 3-D models into SSSCs are still a research topic. There is no existing methodology at this time. We thus undertake a pragmatic, data driven approach to estimating SSSC error surfaces at this time.

In lieu of detailed empirical error maps or error surfaces produced by a good theory of error propagation from uncertainties in the 3-D model, certain simplified summaries of misfit provide the best information we have about errors in predicted travel times. As an example, we plot average 1-norm misfit versus epicentral distance in Figure 1a for all of the stations across Eurasia and for the Group 1 and Group 2 IMS stations and surrogates.

We used the EHB dataset in this analysis and selected events that are considered to be accurate to within 15 km (i.e., GT15 or better) for this analysis. As part of this project, Group2 has developed a raytracer that is capable of predicting accurate travel times for 3-D models at regional distances. We used this raytracer to predict the travel times of ≈ 1,000,000 Pn rays and compared those times to that obtained from the EHB catalog. As shown in Figure 1a, average misfit grows nearly linearly until about 15° and then decreases. This 10° - 15° feature results predominantly from errors in predicting the distance where Pn transitions to a diving-P. This transition is seen in the SSSCs as a high velocity ring that occurs between 10° and 15° at most azimuths. Predictions of the onset of the Pn to diving-P transition are very sensitive to the upper mantle vertical gradient. Therefore, small changes in the model produce moderate changes in travel time predictions in the 10° - 15° range.

We used the standard deviations of the Pn misfits, binned on distance, to obtain model error as a function of distance. Figure 2 shows our model error estimates for Pn along with the estimates currently being used at IDC for Pn. We scaled Pn errors by a factor of two to obtain Sn error estimates. In general, $[\delta Vs/Vs] \approx 2.0\ [\delta Vp/Vp]$ (Robertson and Woodhouse, 1995; extrapolated to upper mantle bottoming depths) and we expect the errors to scale similarly. It is likely that the simply parameterized travel time error surfaces will resemble the overall average curve in Figure 2 scaled to match the observed misfit near each station, thereby reflecting regional variations in modeling error. Average rms-misfits for selected stations with the CUB1.0 model are shown in Figure 3. This figure shows the station coverage of the dataset used for computing error estimates.

3.0 Comparison Between Group2, EHB and IDC Error Estimates

We adopted an azimuthally invariant, region wide model error estimate in Phase 1 that varies only as a function of epicentral distance. This is similar to the baseline currently used by the International Data Center (IDC) in routine location calculations. The IDC model errors for each phase are also only dependent on the source-to-receiver distance. A comparison between these two error estimates is pertinent as Group2 hopes to improve upon the current IDC system. Though the methodologies in deriving the two sets of model errors are similar, the baseline IDC curve was derived using a background 1-D model (i.e., IASP91) while the new errors are based on CUB1.0 model. We expect that with better estimates of the lateral variation of structure, the level of un-modeled signal in our estimates to be smaller than those obtained using IASP91, i.e., we expect our error estimates to be smaller than the baseline at comparable ranges. Figure 4 shows the error estimates for CUB and IASP91. Often times, an analyst can identify the first-arriving P-phase as either P or Pn. If the analyst subsequently uses this as a defining phase in event re-location, the predicted travel times and their associated error estimates can change. To

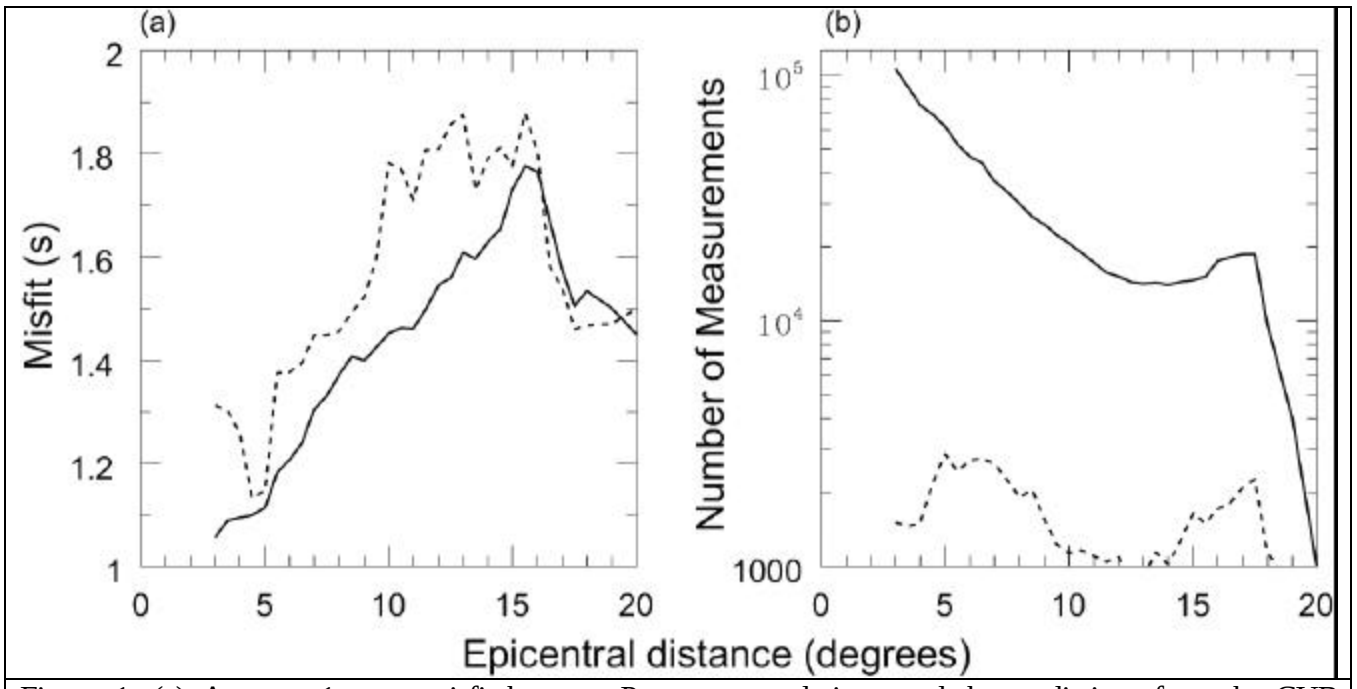


Figure 1. (a) Average 1-norm misfit between Pn wave travel times and the predictions from the CUB 3-D model plotted as a function of epicentral distance. (b) Number of measurements in each 0.5º bin. The solid line is obtained from all of the seismic stations in the Group2 region (Latitude between –15° and 80° and longitude between –40° and 100°) in the EHB catalog and the dashed line is only for the IMS stations in this region. In all, approximately 1,000,000 Pn arrivals were used. The variability of misfit for all stations in this region has been used as a model error estimate.

avoid this ambiguity, we have plotted the baseline IDC error estimates for P, Pn, S and Sn phases. There are few things to note in this comparison:

a. The CUB1.0 error estimates are significantly smaller than the current default estimates.

b. The shapes of the error curves are similar. For instance, for both models, the error estimates increases steadily till about 14º, where there is a significant jump in the values. After about 16º, the error estimates monotonically decrease.

c. The P and the Pn curve are the same for IDC errors at distances less than 17º.

Error estimates for the S-phase at IDC is sparsely sampled. Actually, estimates of errors are only presented at six ranges between 0º - 180º.

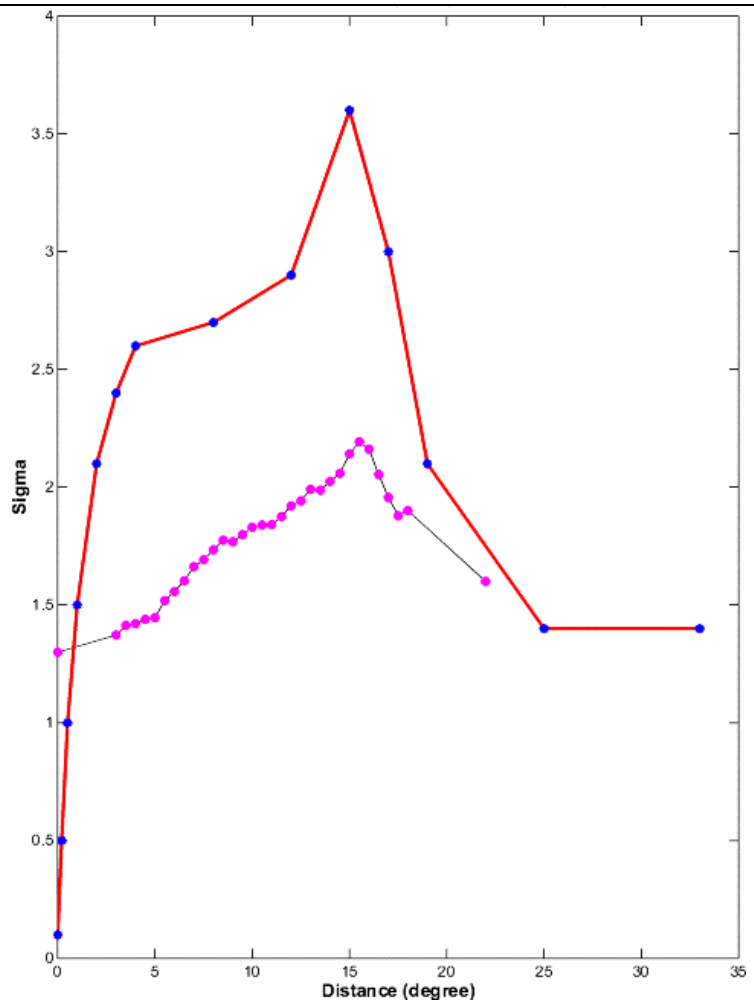
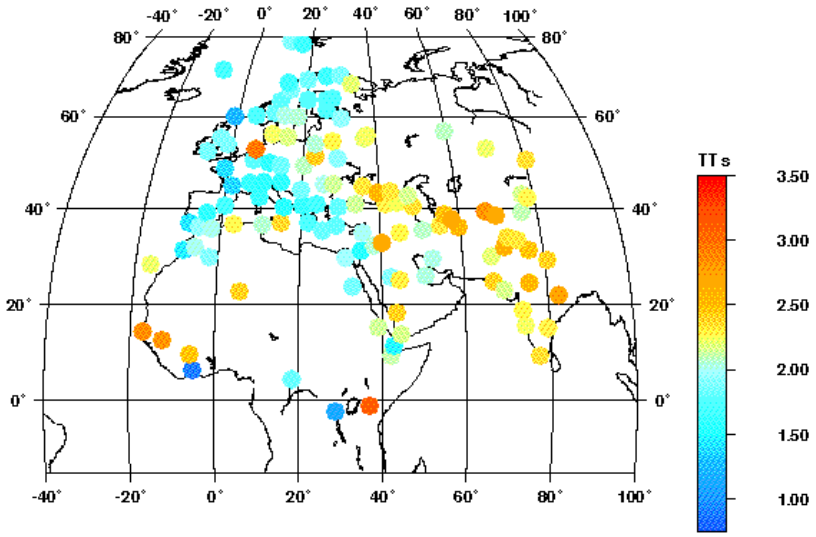


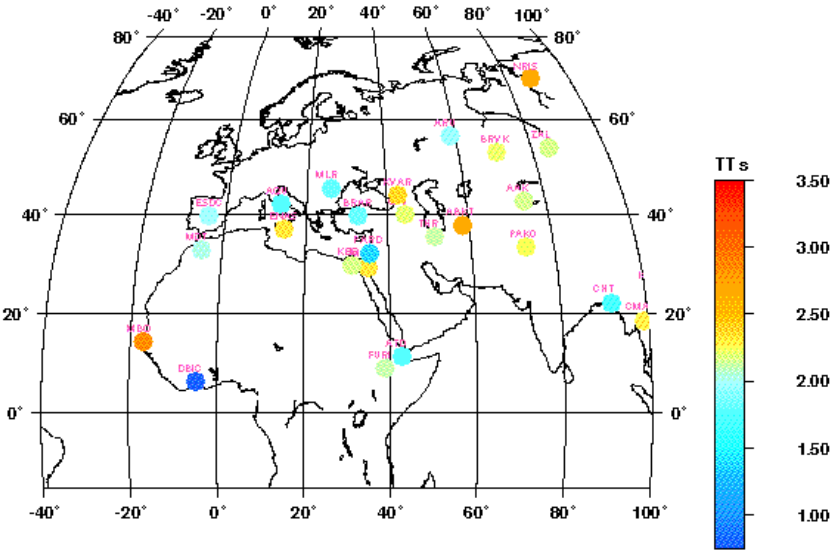
Figure 2. Estimated Pn modeling errors (Phase 1 delivery) is shown by magenta symbols. These errors are only dependent on epicentral distance range and are same for all stations and for both the CUB1.0 and SAIC-HRV models. The current IDC modeling error for Pn/P is shown for reference by the blue symbols. The errors are smoothly interpolated for ranges in between the nodes. The Phase-1 Pn errors are scaled by a factor of two to obtain the Sn errors. Note the similarity in shapes of the two modeling error curves.

Next we compare the error estimates from the high quality EHBGT5 catalog which is a subset of the EHB catalog but only contains events which are candidates for GT5 or better quality. This limits the effect of errors in the source location and origin time on our analysis. We then compute residuals w.r.t. a global 1-D model, *IASP91*. Figure 4 shows our error estimates from the EHBGT5 catalog for Pn and Sn phases. It is interesting to note that for Pn, the EHBGT5 error estimate is similar to the baseline. This is expected as both of them used similar datasets and the same 1-D reference model. On the other hand, the Sn errors differs slightly in amplitude and phase with a prominent peak value at 16°

for the EHBGT5 catalog. This result confirms the general amplitude and distance dependence of modeling errors in current practice at IDC.



(a)



(b)

Figure 3. Summary of rms-misfit between Pn travel time observations and predictions for the CUB1.0 model. (a) Misfits at all of the stations in the Group2 region. (b) Misfits at the IMS stations. We have averaged station-residuals at every 5º, which is comparable to the resolution of the lateral variation of structure for this model (Israelsson et al., 2001).

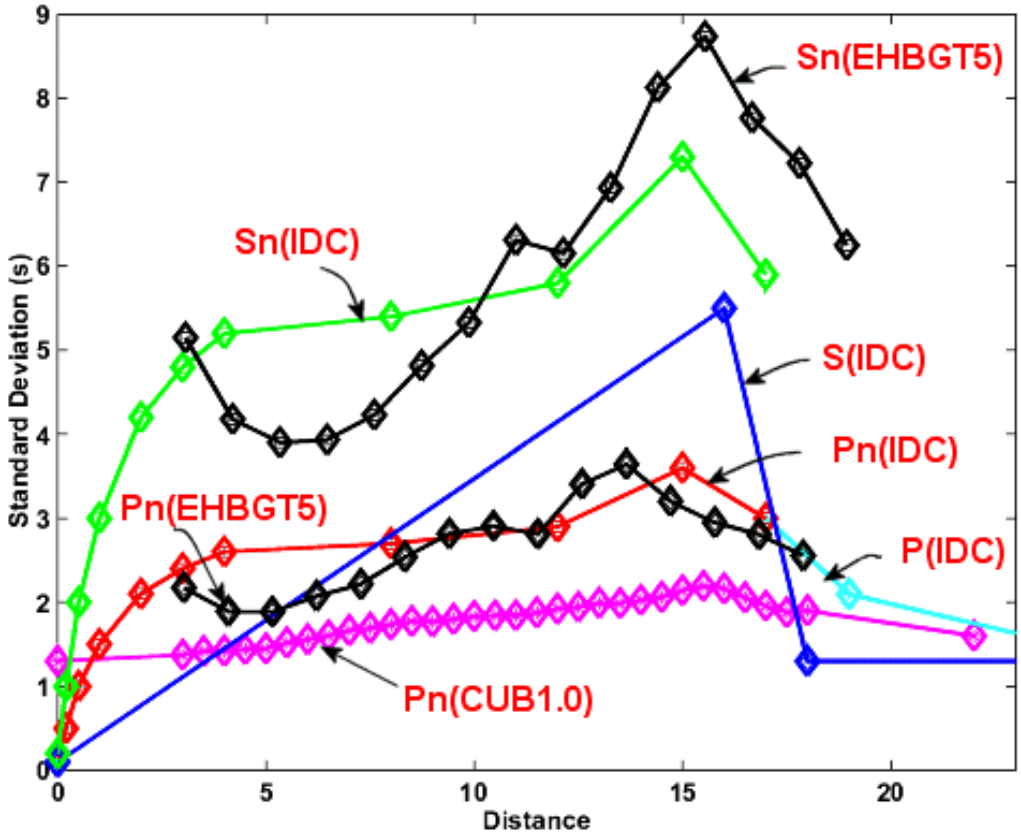


Figure 4. Model error curves for the various datasets analyzed in this study. We show the model error curves used for Phase 1 delivery [Pn(CUB1.0)]. The error curves used in the IDC processing were developed by Israelsson et al. (1997). For the EHBGT5 set, we have used only the events in the EHB catalog that have location accuracy of GT5 or better. The model error is represented as the standard deviation of the residuals w.r.t. IASP91.

4.0 Appraisal of Error Estimates Using Results From Cluster Analysis

In this section, we list the principal findings of the cluster analysis (Israelsson et al., 2001). These findings help us validate the model errors used in Phase 1 and also identify possible future improvements.

- Normalized empirical path corrections obtained from the analysis of each cluster can be used as a high quality travel time dataset, as non-structure related path-

effects on the travel times have been largely eliminated. We can use these datasets, one for each cluster, to test the models by comparing them with model-predicted travel times. In Phase 1, we have developed and carried out validation tests with two separate 3-D models, namely CUB1.0 and SAIC-HRV. Figure 5 shows a comparison of the path effects from the Azgir cluster with the two sets of model predictions. We find that the CUB1.0 model performs significantly better compared to the SAIC-HRV model. Similar results are obtained from the analyses of most other clusters. We can conclude from this test that the CUB model, in general, predicts the travel time dataset better than the SAIC-HRV model and this maybe a reason for the differences in performance in the event relocations using SSSCs for these two models. Secondly, it indicates that the modeling errors probably need to be higher for the SAIC-HRV model.

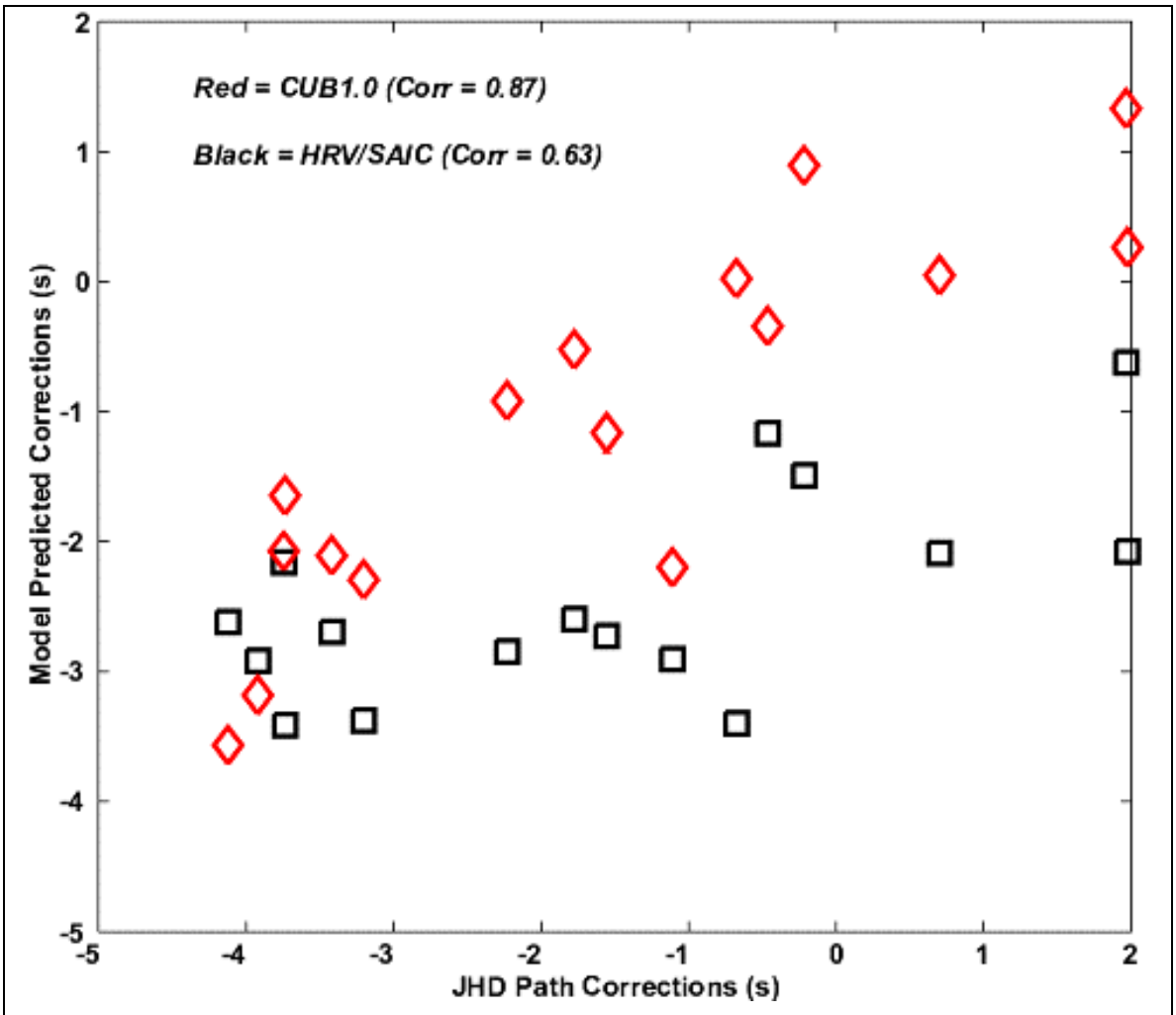


Figure 5. Comparisons of model predictions with empirical path corrections obtained from a JHD analysis (Israelsson et al., 2001) of the Azgir cluster. The path corrections can be used as a high quality dataset of empirical travel times. The CUB1.0 model predicts the empirical data better (corr. coeff. = 0.87 versus 0.63).

- We have computed the correlation coefficients between the JHD corrections and the travel time predictions of the CUB1.0 and SAIC-HRV models for 44 other clusters. A similar analysis for the CUB1.0 model has been carried out by Israelsson et al. (2001). For a detailed discussion of the clusters, please refer to the Israelsson et al. study. Our analyses of the CUB1.0 model differ slightly as in this study we do not remove the outliers before computing the correlation coefficients. As in the Israelsson et al. study, we have used data only up to 18° in distance range and the residuals for each dataset have been normalized by their respective median values to remove the biasing effects of unknown event times on the JHD path corrections. Table 1 shows the results of this correlation analysis that indicates that the correlations are, generally, higher for the CUB1.0 model, i.e., this model predicts the travel time better. The range of travel times for both the model predictions and the path corrections are larger than our modeling error estimates and the difference indicates the level of travel time signals that can be used for event relocations. Surprisingly, the range of values and the standard deviation of path predictions of the SAIC-HRV model are larger than those for the CUB1.0 model though the former is comparatively a smoother model. Figure 6a, b shows the correlation coefficient as a function of cluster location for the CUB1.0 and SAIC-HRV models. Note that there are higher correlation coefficient values for the CUB1.0 model with most of the higher values, for both models, are for clusters to the north of the Group2 region.

- The model error estimates developed in our study has been empirically derived and is most definitely biased upwards by the un-estimated measurement errors. Israelsson (2001) gives a summary of estimates of measurement error based on differences between JHD path corrections that are not significantly biased by travel time model and origin time errors. For a given cluster, the average arrival time differences to a given station pair is compared with travel time difference calculated from the CUB model. This difference represents a residual of the model. With many such residuals, a standard deviation can be calculated for all of subsets of the data that in turn let us estimate measurement errors in the dataset. If we assume that the measurement errors derived from the JHD analysis is representative of the errors in the EHB catalog and that the measurement and modeling errors are independent and their distribution is Gaussian, the difference between their variances can give us estimates of the “true” modeling error. In Figure 7, we show this updated modeling error along with the estimates of our current modeling errors and the measurement errors. We need to test the assumption that modeling error is only distance dependent before we can implement this updated modeling error in the SSSCs.

Cluster Name	Range						STD1	STD2	STD2	CC1	CC2
	JHD		CUB1.0		SAIC-HRV						
	min	max	min	max	min	max					
Algeria, Blida	-2.199	3.252	-2.003	1.157	-2.154	2.276	1.08	0.51	1.09	0.35	-0.47
Algeria, Bordjbou	-2.907	4.841	-1.373	1.044	-2.492	1.469	1.2	0.47	1.02	0.03	-0.17
Algeria, El Asnam	-2.686	7.556	-2.271	1.265	-2.14	3.35	1.59	0.73	1.28	0.14	-0.34
Algeria, Mascara	-1.975	2.608	-2.118	1.822	-2.317	1.935	0.96	0.87	1.31	0.57	-0.69
France, Annecy	-1.605	3.18	-0.968	0.728	-2.614	0.639	1.22	0.47	0.73	0.69	0.45
Greece, Alani	-1.585	5.303	-0.762	0.463	-2.336	1.304	1.32	0.29	0.97	-0.07	0.13
Greece, Amfissa	-2.662	5.638	-2.196	1.113	-8.292	2.118	1.51	0.71	1.9	0.07	-0.15
Greece, Crete	-3.688	7.309	-1.714	0.929	-11.06	1.243	1.73	0.66	1.54	0.19	0.08
Greece, Ionian Sea	-3.144	8.269	-1.526	1.239	-2.268	2.46	1.41	0.65	0.78	0.39	0.39
Greece, Kanallkion	-4.375	10.427	-2.821	1.275	-3.668	2.453	1.57	0.55	1.66	0.42	0.24
Greece, Kefallnia	-5.086	5.505	-2.608	0.936	-9.1	2.432	1.64	0.6	1.84	0.52	0.36
Greece, Pagasae	-3.127	4.864	-1.95	1.802	-3.222	2.553	1.52	0.63	1.76	0.32	0.04
Greece, Thermum	-2.765	5.325	-2.199	1.201	-3.819	2.351	1.73	0.69	1.78	-0.01	-0.23
Greece, Thivai	-3.218	7.356	-1.492	1.624	-3.074	2.702	2.04	0.75	1.69	0.08	-0.29
Greece, Zakynthos	-3.958	8.602	-1.905	1.431	-4.745	1.626	1.99	0.51	1.59	0.01	-0.34
Greece, Zmfissa	-3.217	4.084	-3.221	0.891	-3.356	2.204	1.45	0.63	1.52	0.37	0.08
Israel, Aqaba Central	-3.14	3.227	-1.859	1.247	-2.843	0.849	1.06	0.69	1.08	0.12	-0.16
Israel, Aqaba North	-2.149	7.471	-1.601	1.415	-2.738	1.291	2.01	0.69	1.11	0.12	-0.25
Israel, Aqaba South	-1.455	6.36	-1.905	3.115	-2.835	1.127	1.51	0.81	1.14	0.31	-0.35
Italy, Abruzzo	-2.125	3.6	-0.912	0.699	-2.935	2.412	1.17	0.41	1.19	0.15	-0.1
Italy, Forli	-2.334	2.159	-0.802	0.619	-1.627	1.554	0.89	0.31	0.79	0.03	-0.1
Italy, Gemona	-2.36	6.416	-3.875	1.552	-4.128	2.113	1.22	0.96	1.6	0.04	-0.11
Italy, Reggio	-1.686	2.126	-1.05	1.062	-2.325	0.7	0.85	0.6	0.6	0.35	0.27
Italy, Rionero Central	-4.366	7.38	-2.195	0.802	-3.25	1.982	1.31	0.43	1.36	0.34	0
Italy, Rionero North	-3.325	3.149	-0.877	0.837	-1.865	2.267	1.22	0.4	1.21	0.01	-0.11
Italy, Rionero South	-3.588	14.476	-2.104	0.71	-2.169	2.583	1.99	0.49	1.27	-0.02	-0.2
Italy, Taormina	-3.094	5.617	-0.774	0.722	-20	99	1.63	0.32	49.15	-0.03	0.2
Italy, Umbria-Marche	-3.284	3.261	-1.861	1.011	-4.373	2.46	1.2	0.48	1.29	0.23	-0.2
Italy, Ustica	-2.92	6.057	-0.791	1.471	-106	13	1.74	0.41	45.97	0.36	-0.06
Montenegro, Kotai	-3.057	3.198	-3.071	1.089	-2.892	2.542	1.18	0.68	1.59	0.38	0.15
Morroco, Alhoceima	-1.432	2.365	-1.686	0.786	-3.907	1.43	0.9	0.57	1.33	0.12	-0.01
Morroco, Melilla	-1.931	2.126	-1.354	1.326	-4.099	1.459	0.84	0.57	1.65	0.25	0.09
Poland, Lubin	-5.813	3.663	-3.943	1.009	-3.504	3.654	1.53	0.96	1.39	0.74	0.35
Poland, Silesia	-5.315	1.776	-4.363	1.298	-4.361	2.818	1.82	1.76	2.22	0.8	0.83
Russia, Astrakhan	-2.373	7.368	-1.469	2.15	-0.343	1.811	2.63	0.93	0.6	0.38	0.15
Russia, Azgir	-2.457	3.639	-2.526	2.37	-0.754	2.033	2.04	1.43	0.8	0.87	0.63
Russia, Racha East	-6.725	2.928	-4.698	2.129	-1.551	3.713	2.15	1.68	1.2	0.76	0.27
Russia, Racha West	-3.912	6.131	-4.106	2.71	-1.064	3.608	2.78	1.81	1.44	0.35	0.01
Slovenia, Krm Mountains	-1.761	2.168	-0.941	0.894	-2.02	2.075	0.93	0.44	0.78	-0.08	0.16
Spain, Jayena	-1.977	3.375	-1.341	1.583	-3.426	2.072	1.22	0.69	1.43	0.32	0.53
Spain, Loja	-1.303	3.206	-0.896	0.835	-3.048	1.277	1.1	0.52	0.94	0.19	0.44
Spain, Murcia	-1.524	2.075	-0.665	0.977	-2.209	0.583	0.8	0.42	0.78	0.33	-0.09

Table 1. Results of correlation analysis between empirical travel time (JHD) and model predictions. STD1, STD2 and STD3 indicate the standard deviation of the JHD, CUB1.0 and SAIC-HRV travel times respectively. CC1 and CC2 are the JHD vs. CUB1.0 and JHD vs. SAIC-HRV correlation coefficients.

- In Figure 14 of Israelsson et al. (2001), we compare the standard deviation of the differences between CUB1.0 and JHD path corrections for all of the clusters, as a function of distance, versus the modeling error developed here. The differences reflect the part of the seismic structure not adequately modeled by the CUB1.0 model and thus can be used as a proxy for modeling error. This comparison shows that the modeling errors used by Group2 in this delivery is probably too conservative.

- In Figure 8 of Israelsson et al. (2001), we show that the estimated JHD path corrections at distances less than 1° apart, for both between clusters and between stations, are highly correlated. In general, the correlation length of structure in the Group2 region is about 1° supporting the choice of 1°×1° parameterization of the SSSC surfaces.

- The models estimate structure well at length scale greater than 500 km. An analysis of the relationship between JHD path corrections and CUB1.0 model predictions indicate that they are significantly correlated at distances greater than 5°. This suggests that smaller structures are probably not adequately captured by the models and thus, in Phase 2, we might need to increase the model errors at distances less than 5° from the station.

5.0 Appraisal of Error Estimates Using Results From Event Relocations

Yang et al. (2001) have performed validation testing of the SSSCs, which includes travel time predictions and model errors, by relocating events in the Group2 study area, with and without the SSSCs. Four data sets were used for these tests: a. Fennoscandian events; b. Group2 GT0-GT10 reference events; c. Mid-ocean ridge and transform events, and d. events in the EHB catalog (Engdahl et al., 1998) which are probably of GT5 quality or better. To carry out a comprehensive analysis, the authors used various combinations of calibrated (with SSSCs) and uncalibrated (without SSSCs) regional and uncalibrated teleseismic arrivals. Most of the relocations are from the GT0-GT10 dataset; we discuss the results from those relocations here. Relocation results using SSSCs show overall improvement in event location and confidence ellipses. All events have reduced confidence ellipses without losing 90% coverage. The median reduction in ellipse area is 2360 km^2 (from 4600 to 2240 km^2). The improvement is similar when only calibrated IMS/surrogate regionals are used, but deterioration is larger when only calibrated IMS regionals are used due to limited number/geometry of stations. For events with no more teleseismic phases than 3 times the number of Pn and Sn phases, there is large improvement when located using IMS/surrogate. Using the 571 GT0 – GT10 events (the major part of the testing), authors conclude that:

a. CUB 1.0 model based SSSCs and model errors performed well w.r.t. IASP91 model. Degradation is less than expected from the error model and test data set.

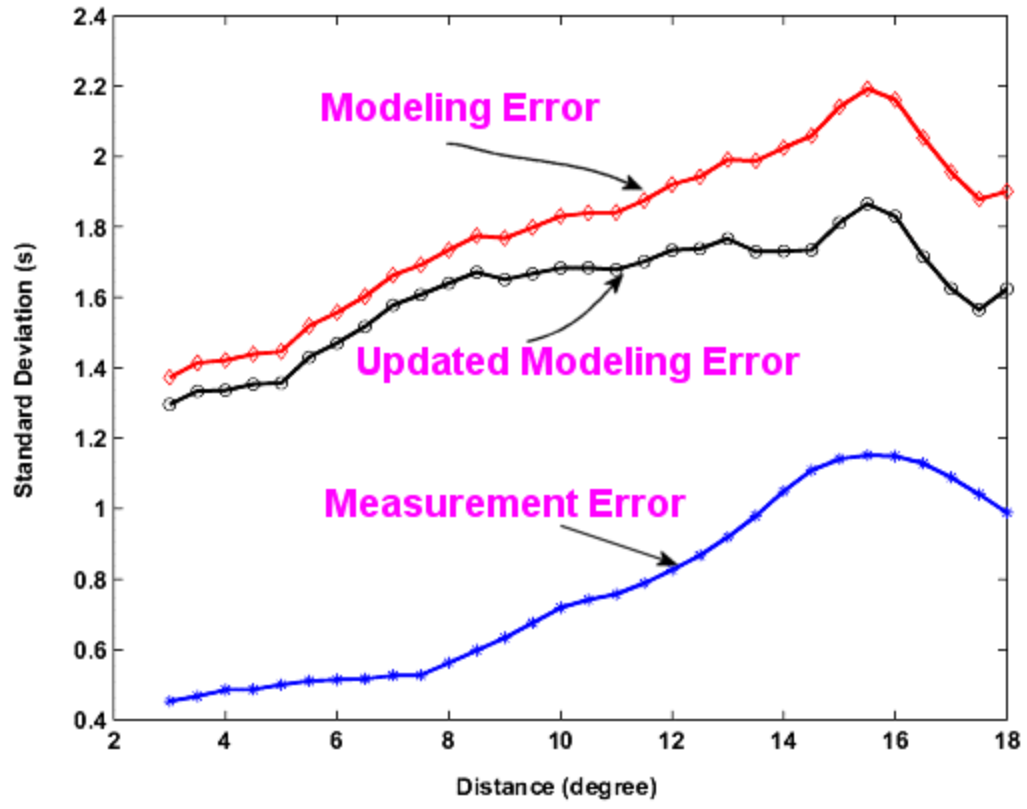


Figure 7. Estimate of distance dependent measurement error as obtained from the JHD analysis. From this limited analysis of the Group2 region, we observe that the measurement error is somewhat correlated with the modeling error and probably contributes to the latter. Thus, robust estimates of measurement error can be used to give us improved estimates of modeling error.

b. Model errors predicted "honest" 90% confidence ellipses.

c. Model errors may be over conservative for 50% of events but under estimated for 5% of the events.

d. IASP91 travel-time tables performed better than should be expected given the current IDC model errors.

Relocation tests using the MORT GT10 events reveal the strong sensitivity of location algorithms to outliers in the distance range of 15 to 20 degrees due to miss association of P to Pn. The percentage of events that failed the 90% coverage test is only slightly below what could be expected based on the sample size. For the validation tests using the candidate GT5 events from the EHB bulletin, 90% coverage reduced from 94% to 84% and the Median coverage is close to the expected value of 0.3. For this dataset, while the model error under predicts at the 90th percentile error, it correctly predicts the

50th percentile coverage. We can thus conclude from the validation testing that our adopted error model is adequate for event relocations in the Group2 region.

McLaughlin and Bondár (2001) investigated the effect on event relocation of uncertainty in reference event location by analyzing the test results of Yang et al. (2001a). The coverage statistic, a measure of GT event location relative to the location confidence ellipse, was computed for event relocations carried out using both calibrated (with SSSCs including our modeling errors) and uncalibrated (without travel time calibrations and using IDC modeling errors) travel times. McLaughlin and Bondár show that both calibrated and uncalibrated travel times perform better than expected for 90% of the events based on the theoretical Chi-squared distribution (Figure 3 of the article). The authors conclude that all values of coverage less than 1, the calibrated model performs measurably better. Importantly, they find that the number of events that degrade is never larger than what we would expect from random chance. These tests validate the use of our modeling errors in the relocations tests using reference events.

6.0 Estimating Model Errors by Analyzing the High Quality Pn/Sn Travel Times Obtained from the IDC REB

The model errors used in this study have been obtained from analysis of the Pn travel times in the EHB catalog. Here we carry out a limited analysis to identify if this dataset is internally consistent. We extracted Pn and Sn arrival times from the PIDC REB till the 2/21/2000 and from the IDC REB for dates after that. We removed arrivals for which SSSCs were used as a change in the baseline velocity model can bias this analysis. Since, we are interested in all of the Pn and Sn arrivals that have been associated by an analyst, we extract data from both defining and non-defining phases. We have used data only from stations within the Group2 region and have events for which there are at least 15 defining phases as these events are deemed to be well located, removing somewhat the bias from location and origin time errors. Finally, we only used events that have been located shallower than 40 km as this is the de facto definition of crustal thickness for an analyst (limits P/Pn identification ambiguity).

First, travel time residuals w.r.t. the IASP91 model are computed. These arrivals are binned as a function of distance and the standard deviation of the residuals within a bin are estimated. The variation of the standard deviation with distance is shown in Figure 8. We note that the curve for Sn is similar to that for the EHBGT5 and the estimated IDC modeling errors (Figure 4). On the other hand, the Pn errors are significantly smaller than the corresponding ones for EHBGT5 and the baseline. This discrepancy between the Pn and Sn datasets is interesting and we present a possible reason for this. To relocate an event in the REB, the analyst makes sure that the residuals for each of the defining phases is less than 2 s (for Pn) and 3 s (for Sn). These residuals are computed for the global 1-D

model IASP91 that might be grossly inadequate for regional phases like Pn and Sn. Thus, there is a tendency for the analyst to pick the arrival time close to the predicted time or make the phase non-defining. The latter factor may be more significant for Sn where only 60% are defining while 85% are defining for Pn. Moreover, there are about four times more Pn defining arrivals than Sn. This suggests that the Pn arrivals are systematically picked to be consistent with the IASP91 predictions thereby yielding smaller standard deviations, compared to Sn, as shown in Figure 8. This experiment suggests that the picked travel time can be a function of the underlying velocity model and thus an use of a more realistic 3-D model, especially in the analysis of regional arrivals, is appropriate.

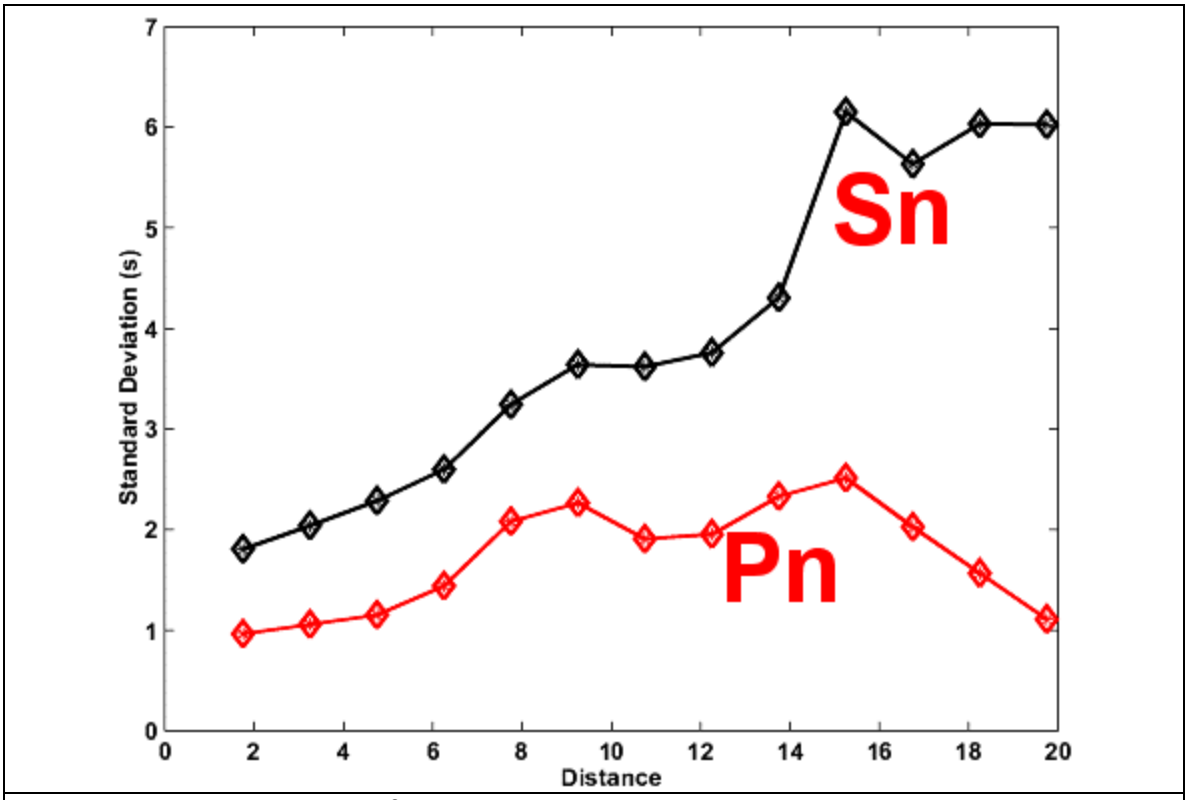


Figure 8. Standard deviation of Pn and Sn residuals, respect to the IASP91 model, obtained from the IDC and PIDC Reviewed Event Bulletin (REB). Please see text for details on data selection.

7.0 Comparison of SSSCs for CUB1.0 and SAIC-HRV Models

One of the components of the modeling error is structure that has not been accounted for in the 3-D models and can bias travel time predictions from SSSCs computed for those models. One of the indicators of this unmodeled signal can be obtained from a comparison of the 3D seismic models of the Group2 region. This can be achieved by comparing the travel time predictions for identical paths for the CUB1.0 and SAIC-HRV

models, i.e., by comparing the corresponding Pn and Sn SSSCs for a station. We carry out this process for the ≈ 750 stations that were used in both the relocation tests using these two models. Since, we used only two separate models in this test, the scope of this test is limited. Moreover, both of these models are optimally smooth and thus error estimates from this test will be conservative. Figure 9 shows a distribution of the difference between these residuals. This figure suggests that the velocities in the SAIC-HRV model, in general, have faster velocities that can contribute to some of the differences observed in the relocation tests with these models (Yang et al., 2001a,b).

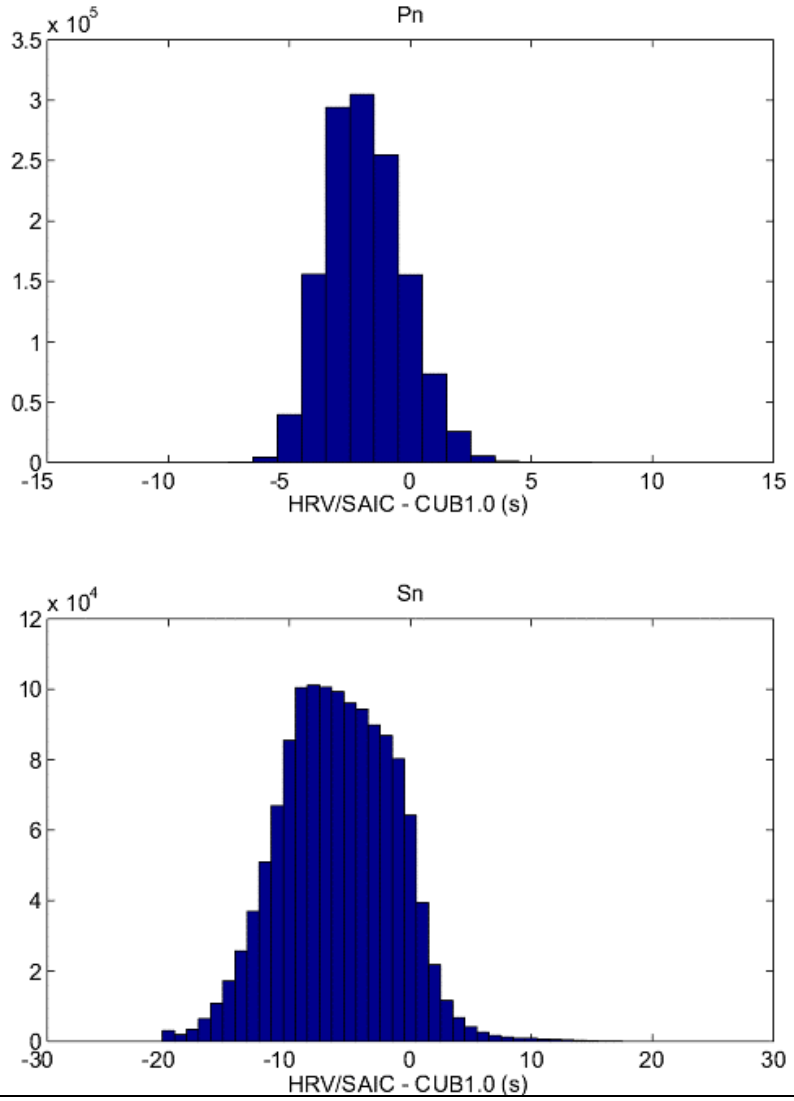


Figure 9. Distribution of the differences of travel time predictions between the SAIC-HRV and CUB1.0 models for Pn and Sn phases. A variation of these residuals as a function of epicentral distance can give us approximates estimates of modeling errors. Note that, in general, we find that the SAIC-HRV is the model with faster velocities.

We show the variation of the travel time differences as a function of epicentral distance in Figure 10. Since, both SAIC-HRV residuals and CUB1.0 residuals have errors

associated with them, we divide the standard deviation of the difference by √2 (assuming that the errors are uncorrelated). Thus, in Figure 10, we show the statistic, σ(SAIC/HRV – CUB1.0)/√2, as a function of epicentral distance. The Pn values are lower than modeling error accepted in the Phase 1 delivery, supporting the hypothesis that the former explains a fraction of the total error budget as the latter also contains the effects of measurement error, origin time errors and the effect of smaller scale structure not modeled by either of these two models.

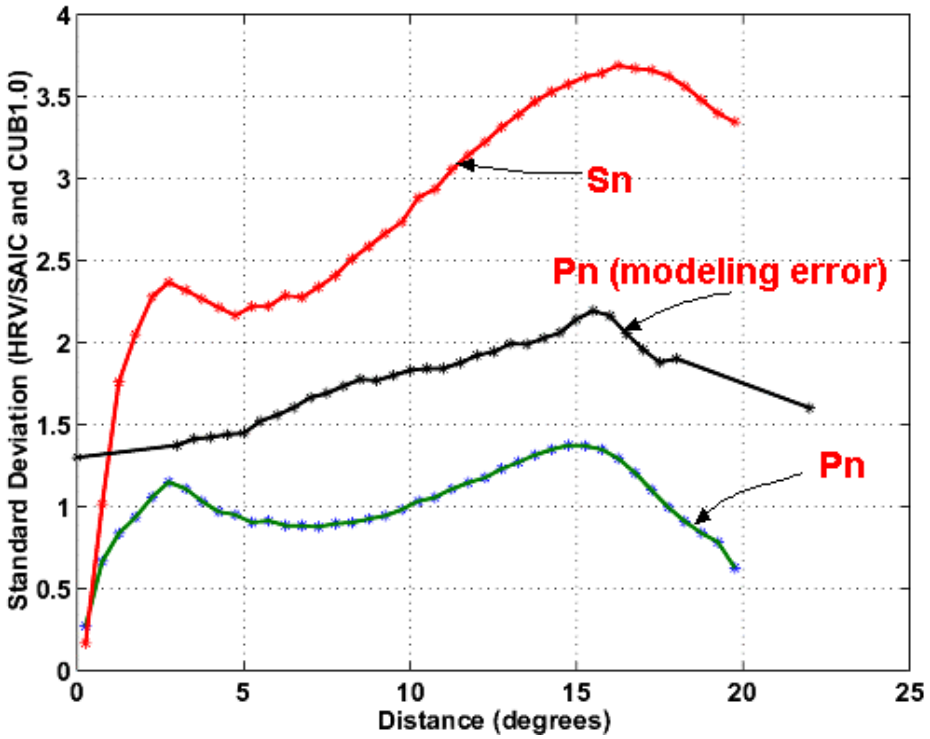


Figure 10. Variation of travel time differences between the SAIC-HRV and CUB1.0 models. We show the variations for the Pn and Sn SSSCs and use the statistic, σ(SAIC/HRV – CUB1.0)/√2, to allow for modeling error in each set of model predictions. Note that they differ by a factor > 2.0. The modeling error adopted in this delivery is shown for reference. The error estimate obtained from the travel time differences is only a part of the error budget and thus, as observed, is a conservative estimate of the modeling error.

8.0 Future Plans: Improving Model Error Estimates in Phase 2

The main goal for computing good estimates of model error is to obtain 90% confidence ellipses that contain 90 % of the reference events. Though this is the case for

nearly all relocation validation tests, our simplified model error estimates need to be improved in Phase 2. The changes will make more realistic estimates, i.e., vary azimuthally as well as radially and be station specific. As described in previous sections, the measurement error can be a significant fraction (up to about 30% in variance) of the modeling error. These improvements will be based both on what we learned in their validation tests in Phase 1 and also from new techniques that will be developed by the Consortium in Phase 2. Initial analyses for some of these techniques have been documented in this report.

1) Model based: Analyzing the variability between SSSC surfaces computed for the two (HRV and CU) models. The SSSCs been derived from 3-dimensional earth models. There are several limitations of these models, e.g., the inability to model seismic structure at short length scales (< 100 km), inadequate spatial sampling both radially and laterally, the use of noisy picks which have significant measurement error and origin time error built into it, etc. Though these limitations exist for each of these models, they vary between models. Thus a comparison between the SSSCs can give us an estimate of the range of model errors. This comparison can be carried out for each station separately and thus individual error surfaces constructed. However, since we are only comparing two separate models, these surfaces will most probably be ill constrained.

2) Data based: Compute 1-D misfit statistics (epicentral distance vs. variance of TT residual) for all of the stations in the region, including non-IMS stations to get good spatial coverage, using a travel time data set and the 3-D model with which the SSSCs are being generated. Using all of the stations will give us a large enough dataset to robustly constrain the region wide variance structure. This is similar to the steps we undertook for Phase 1. In Phase 2, we will improve these estimates by using only GT5 data along with improved Cub and SAIC-HRV models. An initial analysis of this GT5 data (i.e., EHBGT5) using a 1-D model has been presented in this section. There are two possible pathways for this approach: a. We can scale misfits' variances by each station residual (a single number as shown in Figure 3a,b) to estimate station specific, azimuthally invariant error surfaces. (b) For stations with significant number of travel time picks (> 100) we can compute its variogram between pairs of residuals. These variograms can give us estimates of correlation length of structure near the station that can let us extrapolate the modeling errors laterally. We will adopt a simplified, stationary statistics in this analysis.

3) Ad hoc error bounds based on our knowledge of how well we know the structure in a region. This issue can only be addressed after a full-scale comparison of the available 3-D models that will be carried out at the end of Phase 1.

4) The default IASPEI values, maybe in regions of sparse data and/or model coverage. This is a possibility in regions like North Africa where the data coverage, especially for regional body waves, is low. On the other hand, Figure 4

shows that these estimates are considerably high and might not be appropriate for the regional scale relocations that are the focus of this Consortium.

5) Using corrections derived from JHD analysis of event clusters to improve, in limited number of areas, in conjunction with the error surfaces obtained from the above methods. The JHD technique, as described in Israelsson et al., (2001), can let us obtain robust, internally consistent and statistically significant model error estimates in regions close to the event cluster. Unfortunately, the spatial sampling of these clusters is limited. With a significant increase in the number of clusters to be analyzed in Phase 2 along with statistical extrapolation techniques (e.g., variogram based as described in point 2 above) can allow is to use these estimates for region wide, station specific error estimates.

6) Develop a methodology for handling error surfaces where we have reference events. The reference event database (Bondár et al., 2001) is one of the deliverables of Phase 1. We will use this dataset, where some of the ambiguities in event location are removed, to construct improved empirical correction surfaces.

7) Limited Monte Carlo tests with the models. Given the inherent underdetermined nature of the 3-D tomography problem, these tests can allow us to identify the allowable range of model parameters that fits a particular dataset. A robust estimation of this model variability can lead to more realistic estimates of un-modeled signal and thus to improved model error estimate. The University of Colorado group, who is part of this Consortium, is currently exploring this technique.

We expect methods 5 and 7 will be important in estimating model errors in Phase 2.

9.0 Summary and Conclusions

We have chosen a simple and conservative approach to estimate modeling errors for SSSCs delivered in Phase 1. They have been empirically derived using the EHB Pn arrivals by comparing them with the predictions from the CUB1.0 model. Sn model errors have been scaled to Pn. These estimates vary only radially from a station. Moreover, they are the same for both CUB1.0 and SAIC/HRV models for all stations in this region. The model errors are primarily expected to capture the gross uncertainties in travel time as a function of signal only. We have carried out a set of tests to verify if our simplified model errors are valid. At this stage, it is premature to make more sophisticated error models that depend on station and source specificity. We do not expect to develop a formal, error propagation methodology for a high-resolution 3-D seismic model in Phase 2. However, we have carried out an analysis of several estimation techniques that may give more realistic model errors and provide 90% coverage.

From the experiments that we have carried out, we can conclude the following:

1. **Confirmation of IDC baseline errors derived from IASP91 model**

   Comparisons of residuals for the EHBGT5 bulletin, the PIDC and IDC REB's, confirm the general amplitude and distance dependence in current practice for the default IASP91. The 3-D models used in Phase-1 are expected to give better predictions of regional travel times compared to the default 1-D model. The Phase 1 model error estimates are significantly smaller than those currently used at the IDC for the default IASP91 values (Section 3). However, we find that the REB arrivals are picked to be close to the IASP91 predictions and usually within the bounds of the baseline modeling errors (Section 6).

2. **Phase 1 model errors are valid.**

   Model errors, as a function of epicentral distance, is adequate for the following reasons:

a. *More sophisticated error models are not warranted at this time:*

   i. Currently there are no methodology for propagating 3-D model errors to SSSC model errors

   ii. Currently there is no methodology for the evaluation of effect of null spaces in the 3-D models on SSSC model errors.

b. *Empirical model errors obtained from misfit of EHB data with CUB1.0 model can:*

   i. Capture variability in station breakdown (Table A1)

   ii. Capture variability in geographic region (Figure 3)

   iii. Capture variability of un-modeled signal (Section 7)

c. *Empirical (JHD) cluster corrections versus models(Section 4; Israelsson et al., 2001):*

   i. General correlation, especially with the CUB1.0 model predictions, supports the 3-D models

   ii. Scatter of (JHD-CUB1.0) corrections supports overall model errors

   iii. General model error trend supported

   iv. Model error estimates are consistent (Overall, probably smaller than actual model errors, thereby compensating for measurement errors in the model errors.)

d. Evidence from relocation and coverage tests(Section 5; Yang et al., 2001a,b; McLaughlin and Bondár, 2001) :

i. Relocation tests resulted in statistically significant location improvements

ii. Tests using high quality GT0 – GT10 reference events and regional arrivals only all result in 90% or greater coverage, within resolution of sample statistics

iii. One test using lower quality supposed GT5 reference events resulted in 84% coverage

iv. Given the total error budget (model error + measurement error + GT uncertainty), fewer events degraded than would be expected based on random chance

v. In tests simulating the IMS network and using a combination of calibrated Pn and Sn with uncalibrated teleseismics (with/without uncalibrated Pg and Lg), nearly all resulted in 90% or more coverage

vi. Tests using realistic teleseismic-to-regional phase ratios for simulated IMS network verified 90% coverage and location improvement. These tests verified that the relative weighting of teleseismic to regional phases based on our model errors estimates is acceptable

vii. Analysis of test coverage statistics for GT0 – GT10 reference evens indicate that the model errors are conservative at the 50^{th} and 90^{th} percentile

3. Recommendations and future work

This analysis is a report on the progress in the first phase of a two-phase project. In Phase-2, we will present more realistic modeling error estimates more closely tied to the 3-D model and will be a function of geographical location. Towards this goal, we have analyzed and tested several different techniques that can allow us to obtain these improved error estimates. We have identified a JHD based technique and a Monte Carlo model evaluation as most promising for Phase 2 (Section 8). In Section 4, we have shown that the measurement error can be a significant portion of and correlated to modeling error. Thus a separate analysis that focuses on better estimates of estimates of measurement errors will also be needed to improve modeling errors.

10.0 References

Bondár, I. and others, Group-2 Consortium Reference Event List, Group2 report, 2001

Engdahl, E.R., van der Hilst, R.D., and Buland, R.P., 1998, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, Bull. Seism. Soc. Am.,88, 722-743, 1998.

Israelsson, H., Bhattacharyya, J., Bondár, I., McLaughlin, K., and Yang, X., Station Path Corrections based on Event Cluster Analysis, Group2 report, 2001

Israelsson, H., Measurement errors from event clusters, 2001.

Israelsson, H., Swanger, H., and Beall, G., Independent modeling of time measurement and model errors, CMR CCB, 1997.

McLaughlin, K., and Bondár, I., What is the expected deterioration? , Group2 Memo, 2001.

Robertson, G. S., and Woodhouse, J. H., Evidence for proportionality of *P* and *S* heterogeneity in the lower mantle, *Geophys. J. Int.*, 123, 85-116, 1995.

Yang, X., Bondár , I., McLaughlin K., Phase 1 validation test report: Phase 1 testing of the CUB1.0 model, Group2 report, 2001a

Yang, X., Bondár , I., McLaughlin K., Phase 1 validation test report: Phase 1 testing of the SAIC-HRV model, Group2 report, 2001b.